MANCHESTER 1824

The University of Manchester

Structured Nearest Correlation Matrix Problems

Nick Higham
School of Mathematics
The University of Manchester

higham@ma.man.ac.uk
http://www.ma.man.ac.uk/~higham/

Joint work with
Rüdiger Borsdorf, Lijing Lin, Marcos Raydan

Financial Mathematics Day
NAG and Manchester Business School

Numerical Linear Algebra in Finance

- **Correlation matrices: nearness problems.**
- **Roots of transition matrices, $A^{1/p}$.**

Questions From Finance Practitioners

"Given a real symmetric matrix A which is almost a correlation matrix what is the best approximating (in Frobenius norm?) correlation matrix?"

"I am researching ways to make our company's correlation matrix positive semi-definite."

"Currently, I am trying to implement some real options multivariate models in a simulation framework. Therefore, I estimate correlation matrices from inconsistent data set which eventually are non psd."

Correlation Matrix

An $n \times n$ symmetric positive semidefinite matrix A with $a_{ii} \equiv 1$.

Properties:

- symmetric,
- 1s on the diagonal,
- eigenvalues nonnegative,
- off-diagonal elements between -1 and 1.

Correlation Matrix

An $n \times n$ symmetric positive semidefinite matrix A with $a_{ii} \equiv 1$.

Properties:

- symmetric,
- 1s on the diagonal,
- eigenvalues nonnegative,
- off-diagonal elements between -1 and 1.

Is this a correlation matrix?

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Correlation Matrix

An $n \times n$ symmetric positive semidefinite matrix A with $a_{ii} \equiv 1$.

Properties:

- symmetric,
- 1s on the diagonal,
- eigenvalues nonnegative,
- off-diagonal elements between -1 and 1.

Is this a correlation matrix?

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Spectrum: -0.4142, 1.0000, 2.4142.

Stock Research

- Sample correlation matrices constructed from vectors of stock returns.
- Can compute sample correlations of pairs of stocks based on days on which both stocks have data available.
- Resulting matrix of correlations is **approximate**, since built from inconsistent data sets.
- Relatively few vectors of observations available, so approximate correlation matrix has **low rank**.

How to Proceed

- $\surd$ Plug the gaps in the missing data, then compute an exact correlation matrix.
- $\times$ Make ad hoc modifications to matrix: e.g., shift negative e'vals up to zero then diagonally scale.
- $\surd$ Compute the **nearest correlation matrix**.

Problem

Compute distance

$$\gamma(A) = \min\{ \|A - X\| : X \text{ is a correlation matrix} \}$$

and a matrix achieving the distance.

Use a weighted Frobenius norm:

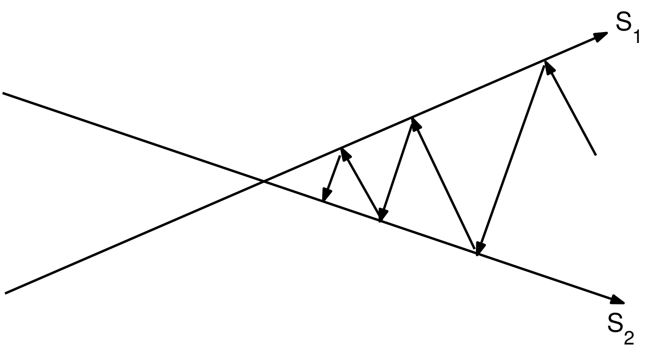
- $\|A\|_W = \|W^{1/2}AW^{1/2}\|_F$ (W pos def),
- $\|A\|_H = \|H \circ A\|_F$ ($h_{ij} > 0$),

where $\|A\|_F^2 = \sum_{i,j} a_{ij}^2$.

★ Constraint set is a closed, convex set, so unique minimizer.

Alternating Projections

von Neumann (1933), for subspaces.



Dykstra (1983) incorporated corrections for closed convex sets.

Projections

For $W \equiv I$.

- For $A = Q\,\mathrm{diag}(\lambda_i)Q^T$ let

$$\mathbf{P_S}(A) := Q\,\mathrm{diag}(\max(\lambda_i, 0))Q^T.$$

- $\mathbf{P_U}(A)$: replace diagonal by 1s.

More complicated for general W; see H (2002).

Algorithm (H, 2002)

Given symmetric $A \in \mathbb{R}^{n\times n}$ this algorithm computes nearest correlation matrix:

$\Delta S_0 = 0,\ Y_0 = A$
for $k = 1, 2, \dots$
  $R_k = Y_{k-1} - \Delta S_{k-1}$ % Dykstra's correction.
  $X_k = \mathbf{P_S}(R_k)$
  $\Delta S_k = X_k - R_k$
  $Y_k = \mathbf{P_U}(X_k)$
end

- X_k and Y_k both converge to solution.
- $O(n^3)$ operations per step.
- Linear convergence.
- Can add further constraints/projections (e.g., Toeplitz).

Newton Method

Qi & Sun (2006): convergent Newton method based on theory of **strongly semismooth matrix functions**.

Denote by A_+ projection of A onto psd matrices.

- Applies Newton to **dual** of $\min \frac{1}{2}\|A - X\|_F^2$ problem:
  $\min_{y \in \mathbb{R}^n} \frac{1}{2}\|(A + \mathrm{diag}(y))_+\|_F^2 - e^T y.$
- Dual problem is ctsly differentiable, but *not twice differentiable* $\Rightarrow$ use generalized Jacobian of gradient.
- **Globally** and **quadratically** convergent.
- Promising test results in Qi & Sun (2006).

The Dual Problem

$$\min_{y\in\mathbb{R}^n} \theta(y) := \frac{1}{2}\|(A+\mathrm{diag}(y))_+\|_F^2 - e^T y.$$

Solution of original problem is $X_* = (A+\mathrm{diag}(y_*))_+$.

Gradient:

$$\nabla\theta(y) = \mathrm{diag}(A+\mathrm{diag}(y))_+ - e.$$

Representative of **generalized Jacobian** of $\nabla\theta(y)$:

$$V_y h = \mathrm{diag}\left(P_y(W_y \circ (P_y^T H P_y))P_y^T\right),$$

where $H = \mathrm{diag}(h)$,

$$A + \mathrm{diag}(y) = P_y \mathrm{diag}(\lambda(y)) P_y^T$$

and

$$W_y = \begin{bmatrix} E_{11} & E_{12} & W_{13} \\ E_{12}^T & 0 & 0 \\ W_{13}^T & 0 & 0 \end{bmatrix}, \qquad E = ee^T.$$

Linear Systems in Newton Method

Newton equation is $V_k d_k = -\nabla\theta(y_k)$ with V_k pos semidef.
Alg requires us to compute d_k s.t.
$\|V_k d_k + \nabla\theta(y_k)\|_2 \le \eta_k \|\nabla\theta(y_k)\|_2$.

- Qi & Sun (2006) used **CG**.
- We use **minres**, which in solving $Ax = b$ **minimizes** $\|Ax_k - b\|_2$ on each iteration, producing monotonically decreasing residuals. Also more suited to semidefinite A than CG.

Hessian V_k

$$V_k h = \text{diag}\left(P_k(W_k \circ (P_k^T H P_k))P_k^T\right), \quad H = \text{diag}(h).$$

- V_k is always positive semidefinite.
- V_k can have "any" spectrum.
- Computing V_k requires $O(n^4)$ flops.

★ For ith column take $H = e_i e_i^T$, $h = e_i$.

★ Get diagonal via $e_i^T(V_k e_i)$.

Preconditioner

Can get **diagonal (Jacobi) preconditioner** in $O(n^3)$ flops:

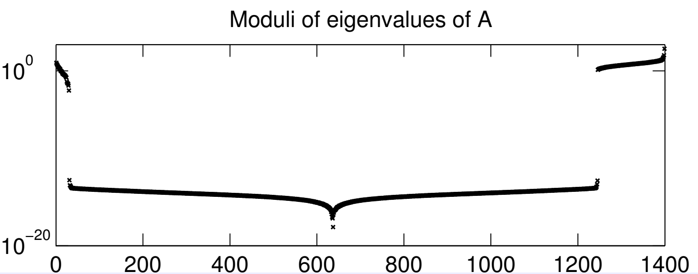
$$\begin{aligned} v_{ii} &= e_i^T P_k (W_k \circ P_k^T e_i e_i^T P_k) P_k^T e_i \\ &= p_i^T (W_k \circ p_i p_i^T) p_i \\ &= p_i^T \mathrm{diag}(p_i) W_k \mathrm{diag}(p_i) p_i \\ &= q_i^T W_k q_i, \end{aligned}$$

where $q_i \in \mathbb{R}^n$ with $q_i = p_i \circ p_i$. Thus

$$\begin{aligned} Q_k &= P_k \circ P_k && n^2 \text{ flops}, \\ M_k &= W_k [q_1 q_2 \dots q_n] = W_k Q_k && \le 2n^3 \text{ flops}, \\ v_{ii} &= q_i^T m_i, \quad i = 1\colon n && 2n^2 \text{ flops}. \end{aligned}$$

Numerical Example from Finance, $n = 1399$

$a_{ii} \equiv 1$, $|a_{ij}| \leq 1$, but not psd. $-8.5 \leq \lambda_i(A) \leq 339$.
A highly rank deficient with 1245 nonpositive ei'vals $\Rightarrow$ $\mathrm{rank}(X) \leq 154$.



$\|A - X_*\|_F = 20.96$.

Numerical Experiment

Stop when $\nabla\theta(y_k) \leq 10^{-7}n$.

cor1399, no precond

	T_{tot}	T_{mvp}	T_{eig}	Iters	# mvp
CG	213	146	63	7	42
Minres	171	104	62	7	30

cor1399, precond

	T_{tot}	T_{mvp}	T_{eig}	Iters	# mvp	T_{pre}
CG	142	77	53	6	22	9
Minres	111	45	53	6	13	9
Alt Proj	529			62		

Other Issues

- Choice of eigensolver: **divide and conquer** is *twice as fast* as **QR** and as **dqdr** on the cor1399 example (NAG MATLAB Toolbox).
- Armijo backtracking rule can be sensitive to rounding errors.
- Computed matrix does not have unit diagonal.
  Solution: set $X \leftarrow D^{-1/2} X D^{-1/2}$ where $D = \mathrm{diag}(X)$.

New Alg Versus Old

tol $= 10^{-7} n$.
Two 387×387 matrices from RiskMetrics.

	nearcor_new		nearcor		Altern. proj.	
	Time	Iter.	Time	Iter.	Time	Iter
cor1399	97	5	378	5	529	62
cor3120	814	4	5256	4	–	–
Risk-daily	0.39	0	0.47	0	1.02	2
Risk-monthly	0.36	0	0.53	0	1.22	2

- New alg is **G02AAF** in NAG Library Mark 22.

Conclusions

- ★ Feasible to compute **nearest** correlation matrix for problems for which spectral decomposition can be done.
- ★ Alternating projections
  - easy to implement,
  - can exploit low rank solutions,
  - linearly convergent,
- ★ Newton method faster and now have robust implementation. Will appear in NAG Library.
- ★ Theory and algorithms for structured problems under development.

References I

L. Anderson, J. Sidenius, and S. Basu.
All your hedges in one basket.
Risk, pages 67–72, Nov. 2003.
|www.risk.net|.

R. Borsdorf.
A Newton algorithm for the nearest correlation matrix.
M.Sc. Thesis, The University of Manchester, Manchester, UK, Sept. 2007.
MIMS EPrint 2008.49, Manchester Institute for Mathematical Sciences, The University of Manchester, UK.

References II

R. Borsdorf and N. J. Higham.
A preconditioned Newton algorithm for the nearest correlation matrix.
MIMS EPrint 2008.50, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Apr. 2008.
Revised November 2008. To appear in IMA J. Numer. Anal.

S. Boyd and L. Xiao.
Least-squares covariance matrix adjustments.
SIAM J. Matrix Anal. Appl., 27(2):532–546, 2005.

References III

I. Grubišić and R. Pietersz.
Efficient rank reduction of correlation matrices.
Linear Algebra Appl., 422:629–653, 2007.

C.-H. Guo and N. J. Higham.
A Schur–Newton method for the matrix *p*th root and its inverse.
SIAM J. Matrix Anal. Appl., 28(3):788–804, 2006.

N. J. Higham.
Computing the nearest correlation matrix—A problem from finance.
IMA J. Numer. Anal., 22(3):329–343, 2002.

References IV

N. J. Higham.
Functions of Matrices: Theory and Computation.
Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

D. K. Knol and J. M. F. ten Berge.
Least-squares approximation of an improper correlation matrix by a proper one.
Psychometrika, 54(1):53–61, 1989.

D. Kurowicka and R. Cooke.
A parametrization of positive definite matrices in terms of partial correlation vines.
Linear Algebra Appl., 372:225–251, 2003.

References V

P. M. Lurie and M. S. Goldberg.
An approximate method for sampling correlated random variables from partially-specified distributions.
Management Science, 44(2):203–218, 1998.

J. Malick.
A dual approach to solve semidefinite least-squares problems.
SIAM J. Matrix Anal. Appl., 26(1):272–284, 2004.

H.-D. Qi and D. Sun.
A quadratically convergent Newton method for computing the nearest correlation matrix.
SIAM J. Matrix Anal. Appl., 28(2):360–385, 2006.

References VI

Z. Zhang and L. Wu.
Optimal low-rank approximation to a correlation matrix.
Linear Algebra Appl., 364:161–187, 2003.